TASK LIST

(We take data set of a restaurant and perform following tasks)

Level 1:

Task 1-Task: Data Exploration and Pre- processing Explore the dataset and identify the number of rows and columns.

Check for missing values in each column and handle them accordingly. Perform data type conversion if necessary.

Analyze the distribution of the target variable ("Aggregate rating") and identify any class imbalances.

Task 2-Task: Descriptive Analysis Calculate basic statistical measures (mean, median, standard deviation, etc.) for numerical columns. Explore the distribution of categorical variables like "Country Code, " "City, " and "Cuisines. " Identify the top cuisines and cities with the highest number of restaurants.

Task 3-Task: Geospatial Analysis Visualize the locations of restaurants on a map using latitude and longitude information. Analyze the distribution of restaurants across different cities or countries. Determine if there is any correlation between the restaurant's location and its rating.

Level-2

Task 1-Task: Table Booking and Online Delivery Determine the percentage of restaurants that offer table booking and online delivery.

Compare the average ratings of restaurants with table booking and those without.

Analyze the availability of online delivery among restaurants with different price ranges.

Task 2-: Price Range Analysis Determine the most common price range among all the restaurants.

Calculate the average rating for each price range. Identify the color that represents the highest average rating among different price ranges.

Task 3- Feature Engineering Extract additional features from the existing columns, such as the length of the restaurant name or address.

Create new features like "Has Table Booking" or "Has Online Delivery" by encoding categorical variables.

Level -3

Task 1-Predictive Modeling

Build a regression model to predict the aggregate rating of a restaurant based on available features.

Split the dataset into training and testing sets and evaluate the model's performance using appropriate metrics.

Experiment with different algorithms (e.g., linear regression, decision trees, random forest) and compare their performance.

Task 2-Customer Preference Analysis

Analyze the relationship between the type of cuisine and the restaurant's rating.

Identify the most popular cuisines among customers based on the number of votes.

Determine if there are any specific cuisines that tend to receive higher ratings.

Task 3- Data Visualization

Create visualizations to represent the distribution of ratings using different charts (histogram, bar plot, etc.).

Compare the average ratings of different cuisines or cities using appropriate visualizations.

Visualize the relationship between various features and the target variable to gain insights.